

日 本 国 特 許 庁
JAPAN PATENT OFFICE

K. Tsukahara

8/29/03

Q77144

10f1

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2002年 8月30日

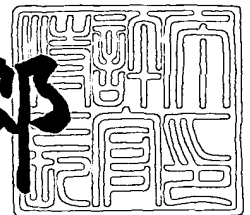
出 願 番 号
Application Number: 特願2002-255388
[ST. 10/C]: [JP2002-255388]

出 願 人
Applicant(s): エヌイーシーコンピュータテクノ株式会社

2003年 7月10日

特許庁長官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3055497

【書類名】 特許願

【整理番号】 03990000

【あて先】 特許庁長官殿

【国際特許分類】 G06F 13/14
G06F 13/10

【発明者】

【住所又は居所】 山梨県甲府市大津町 1 0 8 8 - 3 エヌイーシーコンピ
ュータテクノ株式会社内

【氏名】 塚原 克美

【特許出願人】

【識別番号】 000168285

【氏名又は名称】 エヌイーシーコンピュータテクノ株式会社

【代理人】

【識別番号】 100093595

【弁理士】

【氏名又は名称】 松本 正夫

【手数料の表示】

【予納台帳番号】 057794

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9908108

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 フォールトトレラントコンピュータ及びトランザクション同期制御プログラム

【特許請求の範囲】

【請求項 1】 複数の CPU モジュールで、同じ命令列をクロック同期させて処理するロックステップ方式のフォールトトレラントコンピュータにおいて、
複数の CPU モジュールからの複数かつ同一の I/O トランザクションについて、I/O モジュールのデバイスコントローラ毎に、一定の時間内において互いの一致をチェックし、一致が得られた前記 I/O トランザクションを同期外れではないとみなすことを特徴とするフォールトトレラントコンピュータ。

【請求項 2】 前記 I/O モジュールのデバイスコントローラ毎に、複数の CPU モジュールから発行される前記 I/O トランザクションを入力し、互いの I/O トランザクションのシーケンスの一致を、一定時間待合せて判別するトランザクション同期制御部を備えることを特徴とする請求項 1 に記載のフォールトトレラントコンピュータ。

【請求項 3】 前記トランザクション同期制御部は、シーケンスの一致が得られたタイミングで、前記 I/O トランザクションを前記デバイスコントローラに出力することを特徴とする請求項 2 に記載のフォールトトレラントコンピュータ。

【請求項 4】 前記トランザクション同期制御部は、一定時間内にシーケンスの一致が得られない場合、あるいは I/O トランザクションのシーケンスが互いに相違している場合に、障害の通知を出力することを特徴とする請求項 2 又は請求項 3 に記載のフォールトトレラントコンピュータ。

【請求項 5】 前記トランザクション同期制御部は、
前記複数の CPU モジュールから発行される前記 I/O トランザクションを格納する複数の格納手段と、該格納手段に格納された前記 I/O トランザクションのシーケンスを互いに比較する比較手段と、前記一定時間を計るタイマ手段とを備えることを特徴とする請求項 2 から請求項 4 の何れか 1 項に記載のフォールトトレラントコンピュータ。

【請求項 6】 複数の CPU モジュールで、同じ命令列をクロック同期させて処理するロックステップ方式のフォールトトレラントコンピュータにおいて実行され、

複数の CPU モジュールからの複数かつ同一の I/O トランザクションについて、I/O モジュールのデバイスコントローラ毎に、一定の時間内において互いの一致をチェックし、一致が得られた前記 I/O トランザクションを同期外れではないとみなす処理を行うことを特徴とするフォールトトレラントコンピュータのトランザクション同期制御プログラム。

【請求項 7】 前記 I/O モジュールのデバイスコントローラ毎に、複数の CPU モジュールから発行される前記 I/O トランザクションを入力し、互いの I/O トランザクションのシーケンスの一致を、一定時間待合せて判別する処理を実行することを特徴とする請求項 6 に記載のフォールトトレラントコンピュータのトランザクション同期制御プログラム。

【請求項 8】 シーケンスの一致が得られたタイミングで、前記 I/O トランザクションを前記デバイスコントローラに出力する処理を実行することを特徴とする請求項 7 に記載のフォールトトレラントコンピュータのトランザクション同期制御プログラム。

【請求項 9】 一定時間内にシーケンスの一致が得られない場合、あるいは I/O トランザクションのシーケンスが互いに相違している場合に、障害の通知を出力する処理を実行することを特徴とする請求項 7 又は請求項 8 に記載のフォールトトレラントコンピュータのトランザクション同期制御プログラム。

【請求項 10】 前記複数の CPU モジュールから発行される前記 I/O トランザクションを複数の格納手段に格納する処理と、該格納手段に格納された前記 I/O トランザクションのシーケンスを互いに比較する処理と、前記一定時間の経過を計る処理とを実行することを特徴とする請求項 7 から請求項 9 の何れか 1 項に記載のフォールトトレラントコンピュータのトランザクション同期制御プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数のCPUモジュールで、同じ命令列をクロック同期させて全く同じに処理するロックステップ方式のフォールトトレラントコンピュータ (fault tolerant computer) に関し、特に、複数のCPUモジュールからIOコントローラに対して発行されるIOトランザクションの時間的なずれを許容するフォールトトレラントコンピュータ及びトランザクション同期制御プログラムに関する。

【0002】**【従来の技術】**

従来のハードウェアを多重化し同一クロックで同期するフォールトトレラントコンピュータは、複数のCPUモジュールと複数のIOモジュールを有し、各IOモジュールは複数のCPUモジュールの同期が一致しているかどうかをチェックする比較装置をそれぞれ有しており、同時にIOトランザクションを入力した場合にCPUモジュールが同期していると判断してIO処理を行い、各CPUモジュールが故障しない限り同期していることを保証していた。

【0003】

しかし、近年のプロセッサの高速化に伴い、各CPUモジュールに搭載されるプロセッサは、同一のクロックを入力しても完全に同期しない同期外れが発生し、これにより各IOモジュールで複数のCPUモジュールから発行されるIOトランザクション相互に時間的な差分が発生する場合がある。

【0004】

このようなCPUモジュールの障害に起因しない同期外れの現象に関しては、CPUモジュールでのハードウェア動作に差異があるものの、プログラム上は同じ動作がなされていることが分かっている。しかし、従来のフォールトトレラントコンピュータでは、障害が原因の同期外れでないにも関わらず、CPUモジュールの縮退状態や再組み込みの動作が発生してしまうことがある。

【0005】**【発明が解決しようとする課題】**

上述したように、従来のフォールトトレラントコンピュータでは、各CPUモジュールの同期外れにより、各IOモジュールで複数のCPUモジュールから発行

される I O トランザクション相互に時間的な差分が発生した場合、障害が原因の同期外れでないにも関わらず、C P U モジュールの縮退状態や再組み込みの動作が発生してしまうという問題があった。このことは、またフォールトトレラントコンピュータの MTBF (Mean Time Between Failure: コンピュータシステムが故障してから次に故障するまでの平均時間) を小さくする原因となり、フォールトトレラントコンピュータ本来の利点が損なわれてしまう結果ともなる。

【0006】

なお、運用系プロセッサと待機系プロセッサ間で、プロセッサ間のバスのデータ転送開始及び終了の信号を用いて、待機系プロセッサの処理の遅れを待つことにより、短時間で両者の同期をとるといった（例えば、特開平 11-339932 号公報）、複数のプロセッサ間の同期外れを解決するための従来技術については種々提案されているが、何れも I O コントローラ毎の I O トランザクションの時間的なずれを許容することを特徴とする本発明とは相違する。

【0007】

本発明の目的は、時間的にずれて I O トランザクションが発行されても、論理的なシーケンスの一致が得られれば障害が原因の同期外れではないとみなすことにより、障害が原因の同期外れでないにも関わらず C P U モジュールの縮退や再組み込み動作が発生してしまうのを極力抑え、一定範囲のプロセッサの同期外れを許容することができるフォールトトレラントコンピュータ及びトランザクション同期制御プログラムを提案することにある。

【0008】

【課題を解決するための手段】

上記目的を達成する本発明は、複数の C P U モジュールで、同じ命令列をクロック同期させて処理するロックステップ方式のフォールトトレラントコンピュータにおいて、複数の C P U モジュールからの複数かつ同一の I O トランザクションについて、I O モジュールのデバイスコントローラ毎に、一定の時間内において互いの一致をチェックし、一致が得られた前記 I O トランザクションを同期外れではないとみなすことを特徴とする。

【0009】

請求項2の本発明のフォールトトレラントコンピュータは、前記I Oモジュールのデバイスコントローラ毎に、複数のC P Uモジュールから発行される前記I Oトランザクションを入力し、互いのI Oトランザクションのシーケンスの一致を、一定時間待合せて判別するトランザクション同期制御部を備えることを特徴とする。

【0010】

請求項3の本発明のフォールトトレラントコンピュータは、前記トランザクション同期制御部は、シーケンスの一致が得られたタイミングで、前記I Oトランザクションを前記デバイスコントローラに出力することを特徴とする。

【0011】

請求項4の本発明のフォールトトレラントコンピュータは、前記トランザクション同期制御部は、一定時間内にシーケンスの一致が得られない場合、あるいはI Oトランザクションのシーケンスが互いに相違している場合に、障害の通知を出力することを特徴とする。

【0012】

請求項5の本発明のフォールトトレラントコンピュータは、前記トランザクション同期制御部は、前記複数のC P Uモジュールから発行される前記I Oトランザクションを格納する複数の格納手段と、該格納手段に格納された前記I Oトランザクションのシーケンスを互いに比較する比較手段と、前記一定時間を計るタイマ手段とを備えることを特徴とする。

【0013】

請求項6の本発明によるフォールトトレラントコンピュータのトランザクション同期制御プログラムは、複数のC P Uモジュールで、同じ命令列をクロック同期させて処理するロックステップ方式のフォールトトレラントコンピュータにおいて実行され、複数のC P Uモジュールからの複数かつ同一のI Oトランザクションについて、I Oモジュールのデバイスコントローラ毎に、一定の時間内において互いの一致をチェックし、一致が得られた前記I Oトランザクションを同期外れではないとみなす処理を行うことを特徴とする。

【0014】

請求項7の本発明によるフォールトトレラントコンピュータのトランザクション同期制御プログラムは、前記I/Oモジュールのデバイスコントローラ毎に、複数のCPUモジュールから発行される前記I/Oトランザクションを入力し、互いのI/Oトランザクションのシーケンスの一致を、一定時間待合せて判別する処理を実行することを特徴とする。

【0015】

請求項8の本発明によるフォールトトレラントコンピュータのトランザクション同期制御プログラムは、シーケンスの一致が得られたタイミングで、前記I/Oトランザクションを前記デバイスコントローラに出力する処理を実行することを特徴とする。

【0016】

請求項9の本発明によるフォールトトレラントコンピュータのトランザクション同期制御プログラムは、一定時間内にシーケンスの一致が得られない場合、あるいはI/Oトランザクションのシーケンスが互いに相違している場合に、障害の通知を出力する処理を実行することを特徴とする。

【0017】

請求項10の本発明によるフォールトトレラントコンピュータのトランザクション同期制御プログラムは、前記複数のCPUモジュールから発行される前記I/Oトランザクションを複数の格納手段に格納する処理と、該格納手段に格納された前記I/Oトランザクションのシーケンスを互いに比較する処理と、前記一定時間の経過を計る処理とを実行することを特徴とする。

【0018】

【発明の実施の形態】

以下、本発明の実施の形態について図面を参照して詳細に説明する。図1は、本発明の実施の形態によるフォールトトレラントコンピュータの構成を示すブロック図である。

【0019】

図1を参照すると、本実施の形態によるフォールトトレラントコンピュータ10は、複数のCPUモジュール20、30を備えており、各CPUモジュール20

、30はそれぞれクロック同期して同じ命令列を処理し、各CPUモジュールの処理結果を比較し、1個のCPUモジュールで故障が発生しても、残りのCPUモジュールにより処理を継続することができるようになっている。このCPUモジュール20、30は、それぞれ1以上のプロセッサと、プロセッサ外部バスと、メモリを有している。なお、図1には、本実施の形態の特徴的な構成のみを記載し、他の一般的な構成については記載を省略してある。

【0020】

また、フォルトトレラントコンピュータ10は、第1IOモジュール50と第2IOモジュール60を備え、それらの第1IOモジュール50と第2IOモジュール60が上記CPUモジュール20、30の双方に接続されている。

【0021】

この第1及び第2CPUモジュール20、30は、互いに同じクロックで動作し、第1及び第2IOモジュール50、60に対してそれぞれIOTランザクションの発行を行なう。

【0022】

第1及び第2IOモジュール50、60は、それぞれ2つのIOコントローラ52A、52Bと62A、62Bを備え、かつ各IOコントローラ毎にトランザクション同期制御部51A、51Bと61A、61Bを備えて構成される。

【0023】

IOコントローラ52A、52Bと62A、62Bには、例えばネットワークコントローラやディスクコントローラ等を含むデバイスが接続され、これらのデバイスに対してIOTランザクションに従って入出力制御処理を実行する。

【0024】

トランザクション同期制御部51A、51B、61A、61Bは、第1及び第2CPUモジュール20、30からのIOTランザクションについて、そのシーケンスが一致しているかどうかを対応するIOコントローラ毎にチェックし、時間的にずれが発生しても同じIOTランザクションが異なるタイミングで入力されるのを予め設定したい一定時間だけ待ち合わせ、IOTランザクションのシーケンスが一致したタイミングでデバイスに対応したIOコントローラ52A、52

B、62A、62Bに対して出力する。

【0025】

図4に、トランザクション同期制御部51Aの構成例を示す。ここでは、便宜上トランザクション同期制御部51について示しているが、その他のトランザクション同期制御部についても全く同じ構成である。

【0026】

図示のように、トランザクション同期制御部51Aは、第1CPUモジュール20と第2のCPUモジュール30から入力するI/Oトランザクションをそれぞれ格納する第1トランザクション格納手段81-1及び第2トランザクション格納手段81-2と、第1及び第2トランザクション格納手段81-1、81-2双方のI/Oトランザクションのシーケンスの一致を比較して判別する比較手段82と、一定時間の経過を計るタイマ手段83と、比較手段82によるI/Oトランザクションのシーケンスの一致が得られたタイミングでI/OトランザクションをI/Oコントローラに対して出力する出力制御手段84を有する。出力制御手段84は、タイマ手段83による一定時間経過してもI/Oトランザクションのシーケンスの一致が得られない場合には、障害監視部等に障害を通知する。

【0027】

なお、図4に示したトランザクション同期制御部の構成例については、あくまでその一構成例を示すものであり、図示の構成に限定されるものではない。

【0028】

図5は、トランザクション同期制御部の動作内容を説明するフローチャートである。まず、第1CPUモジュール20と第2のCPUモジュール30からのI/Oトランザクションを入力し（ステップ501）、第1及び第2のCPUモジュールのI/Oトランザクションのシーケンスが互いに一致するかどうかを判別する（ステップ502）。

【0029】

I/Oトランザクションのシーケンスの一致していれば、入力したI/OトランザクションをI/Oコントローラへ出力する（ステップ503）。

【0030】

ＩＯトランザクションのシーケンスが一致していない場合には、一定時間が経過しているかを判別し（ステップ５０４）、一定時間が経過していなければＩＯトランザクションの入力を待合せる。一定時間が経過していれば、フォールトレラントコンピュータ内に備えられている障害監視部等に障害の発生を通知する。

【００３１】

次いで、図１、図２、図３を参照して、第１及び第２ＣＰＵモジュール２０、３０からどのようにＩＯトランザクションが発行され、第１及び第２ＩＯモジュール５０、６０でどのように一致チェックが行なわれ、ＩＯトランザクションが処理されるかについて具体例をあげて説明する。

【００３２】

ここで、プロセッサ同期外れの場合、第１ＣＰＵモジュール２０から第１ＩＯモジュール５０へのＩＯトランザクションが、ＩＯトランザクション（ＩＯ－Ａ１）、ＩＯトランザクション（ＩＯ－Ｂ１）、ＩＯトランザクション（ＩＯ－Ａ２）、ＩＯトランザクション（ＩＯ－Ｂ２）の順で発行されたとする。

【００３３】

なお、ＩＯトランザクション（ＩＯ－Ａ１）、（ＩＯ－Ａ２）はＩＯコントローラ５２Ａに対して発行され、ＩＯトランザクション（ＩＯ－Ｂ１）、（ＩＯ－Ｂ２）はＩＯコントローラ５２Ｂに対して発行されるＩＯトランザクションであるとする。

【００３４】

一方、第２ＣＰＵモジュール３０から第１ＩＯモジュール５０へのＩＯトランザクションについては、ＩＯトランザクション（ＩＯ－Ａ１）、ＩＯトランザクション（ＩＯ－Ａ２）、ＩＯトランザクション（ＩＯ－Ｂ１）、ＩＯトランザクション（ＩＯ－Ｂ２）の順番で発行されるたとする。

【００３５】

このＩＯトランザクションの順番を比較すると、図２に示すように、第１及び第２ＣＰＵモジュール２０、３０から第１及び第２ＩＯモジュール５０、６０に対して発行されるＩＯトランザクションの中で、ＩＯトランザクション（ＩＯ－

B1) と (IO-A2) の順番が互いに逆となり、2つのIOトランザクションが不一致となっていることが分かる。

【0036】

第1 IOモジュール50は、通常は同じタイミングでIOトランザクションに差分があるため、通常であればどちらかのCPUモジュールを縮退もしくは再組み込みする動作が行われる状況であっても、IOコントローラ毎にIOトランザクションのシーケンス(順番)をチェックするためのトランザクション同期制御部51A、51Bを備えることで、タイミングがずれても同じIOトランザクションと認識して処理を継続する。

【0037】

例えば、IOコントローラ52Aに対するIOトランザクションをIO-A_xとし、IOコントローラ52Bに対するIOトランザクションをIO-B_xとすると、図3に示すように、トランザクション同期制御部51Aで見ると、第1及び第2CPUモジュールから発行されるIOトランザクションは(IO-A1) - (IO-A2) となって共に一致し、トランザクション同期制御部51Bで見ても第1及び第2CPUモジュールから発行されるIOトランザクションは(IO-B1) - (IO-B2) で一致する。

【0038】

このときトランザクション同期制御部51A、51Bは、第1CPUモジュール20と第2CPUモジュール30との時間的差分が発生する可能性が高いため、IOコントローラ毎にシーケンスの一致をチェックするのみで時間的な一致については判断していない。すなわち、トランザクション同期制御部51A、51Bは、それぞれ第1CPUモジュール20と第2CPUモジュール30からの同じIOトランザクションが到着するまで一定時間待ち合わせした後にIOコントローラ52A及び52Bに対してIOトランザクションを出力する。

【0039】

IOコントローラ52Aに着目すれば、まず最初のタイミングで第1CPUモジュール20と第2CPUモジュール30からの同じIOトランザクション(IO-A1)が発行され、次のタイミングでは第2CPUモジュール30からの(

IO-A2)は届くが、第1CPUモジュール20からのIOトランザクション(IO-A2)が届かない。IOコントローラ52Aに対応するトランザクション同期制御部51Aは、第1CPUモジュール20からのIOトランザクション(IO-A2)が受け取るまで一定時間待合せ(IOトランザクション(IO-A1)、(IO-A2)を出力せずに保持し)、さらに次のタイミングで第1CPUモジュール20からのIOトランザクション(IO-A2)が届くと、IOトランザクションのシーケンスが一致したことを判別し、それらのIOトランザクション(IO-A1)、(IO-A2)を、IOコントローラ52Aに対して出力する。

【0040】

なお、トランザクション同期制御部51A、51B、61A、61Bよる比較の結果、IOコントローラに対する第1、第2CPUモジュール20、30からのIOトランザクションのシーケンスが互いに相違している場合には、フォールトレラントコンピュータ内に備えられている障害監視部等に障害の発生を通知するようにしてもよい。

【0041】

本発明のフォールトレラントコンピュータでは、トランザクション同期制御部が実行する各手段の機能をハードウェア的に実現することは勿論として、上記した各手段の機能を実行するトランザクション制御プログラム100をIOコントローラ等のプロセッサ(CPU)にロードして実行することにより、ソフトウェア的に実現することができる。このトランザクション制御プログラム100は、磁気ディスク、半導体メモリその他の記録媒体に格納され、その記録媒体からCPUのメモリにロードされ、CPUで実行されることにより、上述した各機能を実現する。

【0042】

以上好ましい実施の形態をあげて本発明を説明したが、本発明は必ずしも上記実施の形態に限定されるものではなく、その技術的思想の範囲内において様々に変形して実施することができる。

【0043】

【発明の効果】

以上説明したように本発明によれば、第1及び第2CPUモジュールのプロセッサに同期外れが発生しIOモジュールのIOコントローラに対して同じタイミングでIOトランザクションが発行されなくても、IOコントローラ毎にIOトランザクションのシーケンス一致をチェックし、一定時間内にシーケンスの一致が得られれば障害が原因の同期外れではないとみなして、IOトランザクションをIOコントローラに出力するようにしたので、障害が原因の同期外れでないにも関わらずCPUモジュールの縮退や再組み込み動作が発生してしまうのを極力抑え、一定範囲のプロセッサの同期外れを許容することができる。

【図面の簡単な説明】

【図1】 本発明の実施の形態によるフォルトトレラントコンピュータの構成を示すブロック図である。

【図2】 本発明の実施の形態によるフォルトトレラントコンピュータの各CPUモジュールから発行されるIOトランザクションの具体例を説明する図である。

【図3】 本発明の実施の形態によるフォルトトレラントコンピュータの具体的な動作例を説明する図である。

【図4】 本発明の実施の形態によるフォルトトレラントコンピュータに備えられるトランザクション同期制御部の構成例を示すブロック図である。

【図5】 本発明の実施の形態によるフォルトトレラントコンピュータに備えられるトランザクション同期制御部の動作例を説明するフローチャートである。

【符号の説明】

10 フォルトトレラントコンピュータ

20 第1CPUモジュール

30 第2CPUモジュール

50 第1IOモジュール

60 第2IOモジュール

51A、51B、61A、61B トランザクション同期制御部

5 2 A、5 2 B、6 2 A、6 2 B I Oコントローラ

8 1 - 1 第 1 トランザクション格納手段

8 1 - 2 第 2 トランザクション格納手段

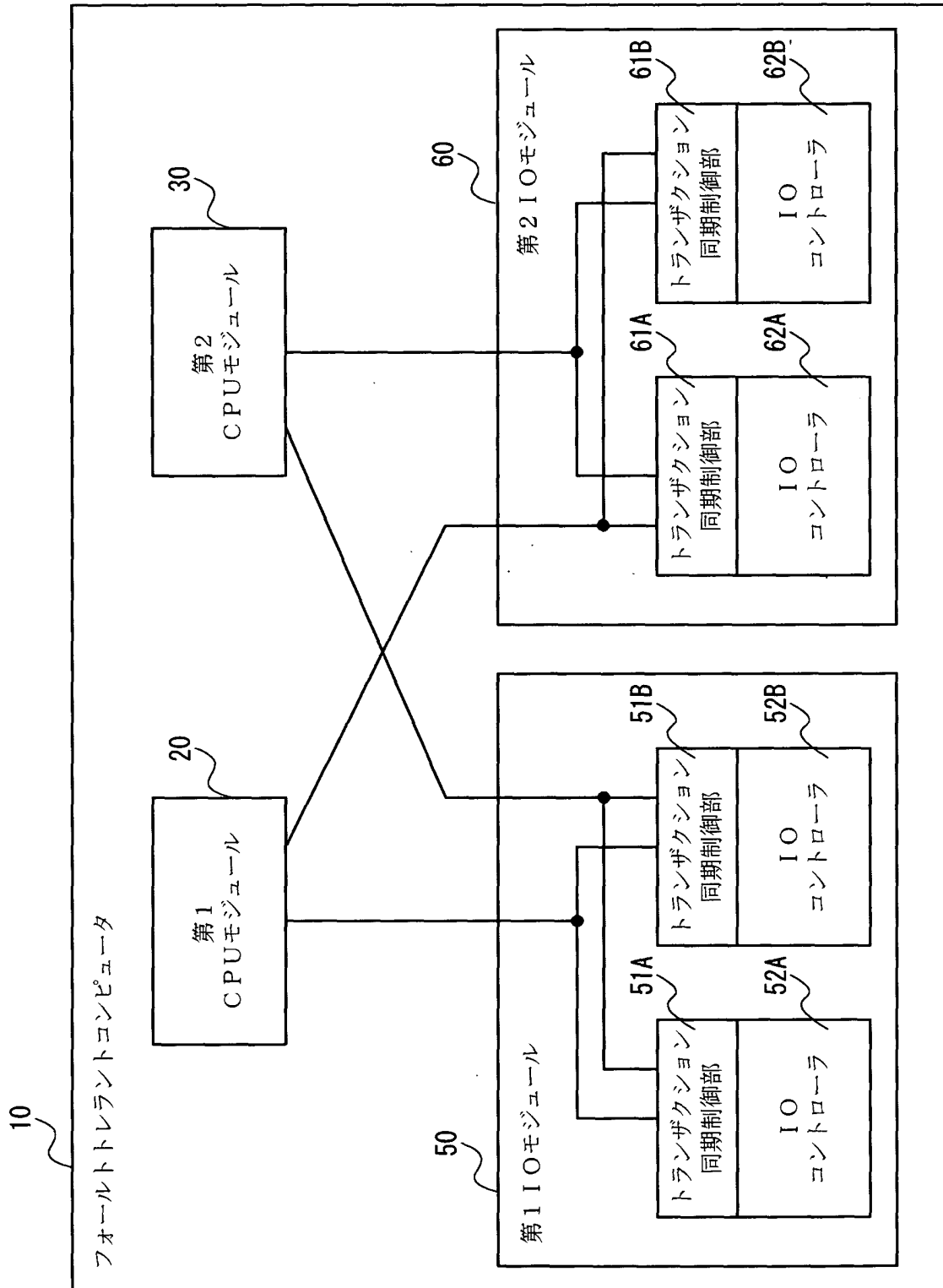
8 2 比較手段

8 3 タイマ手段

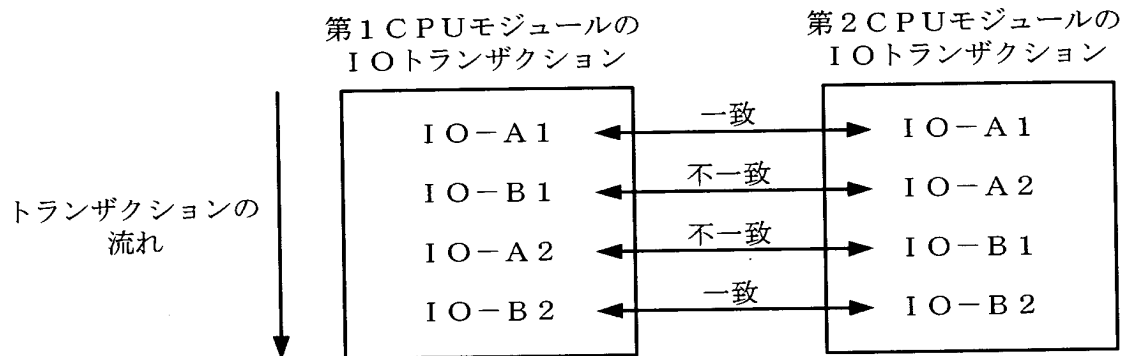
8 4 出力制御手段

【書類名】 図面

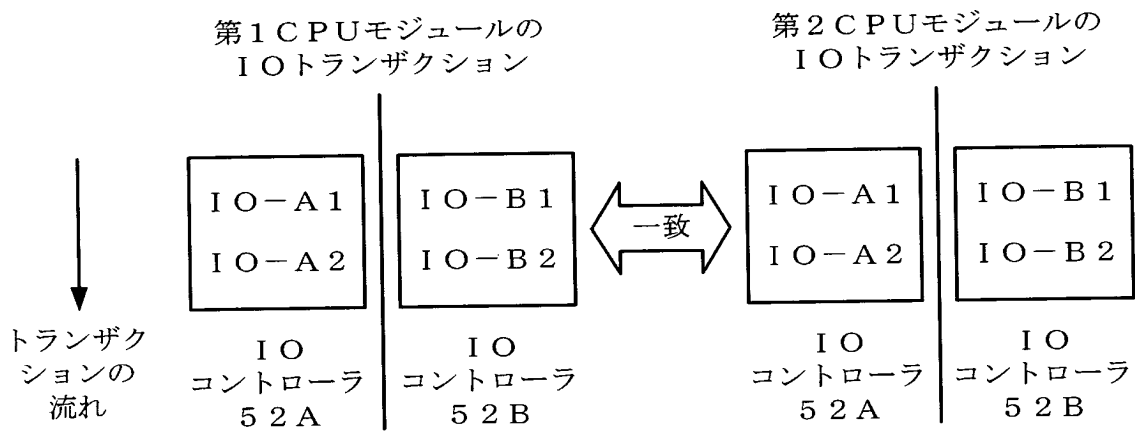
【図 1】



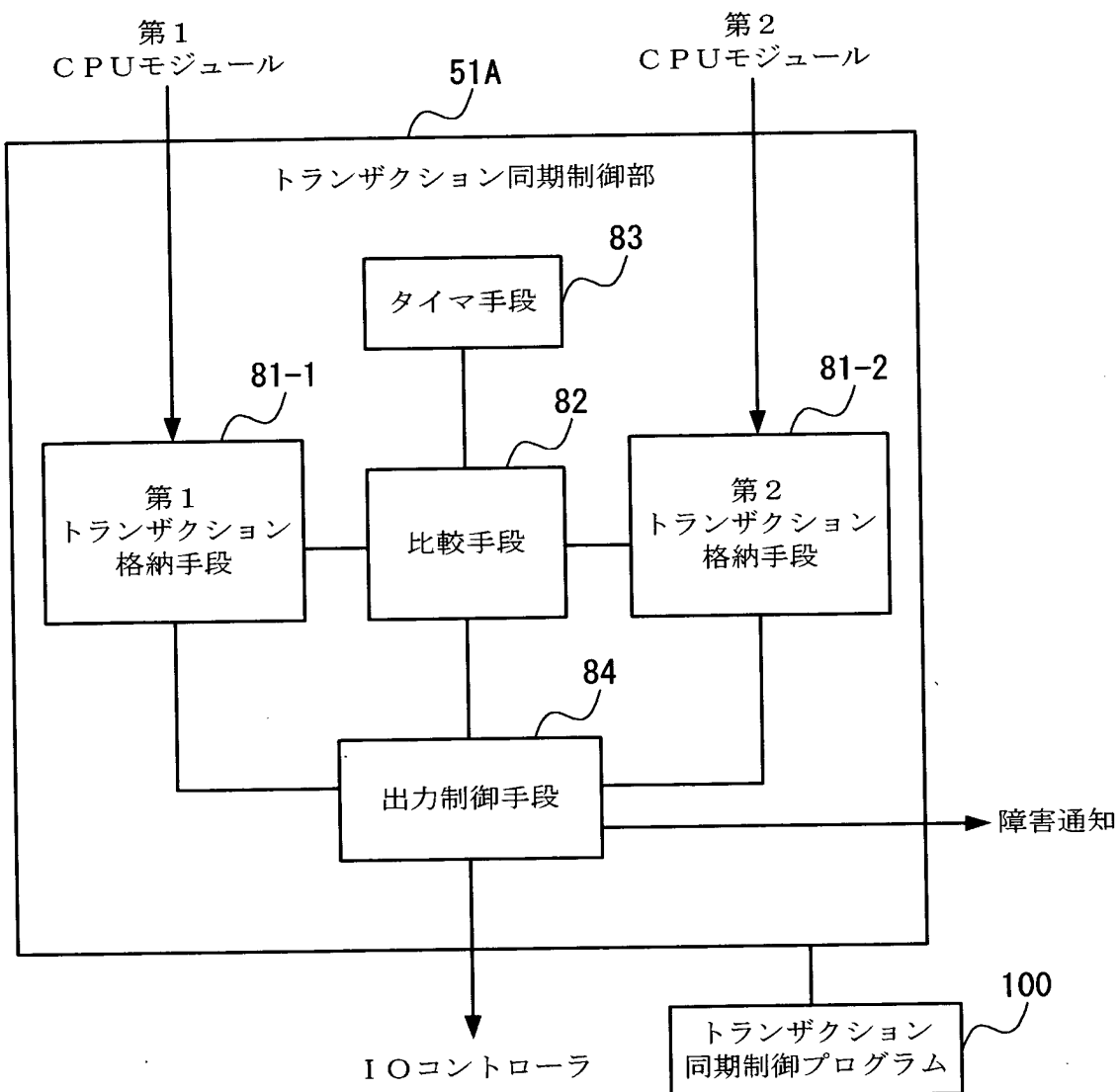
【図 2】



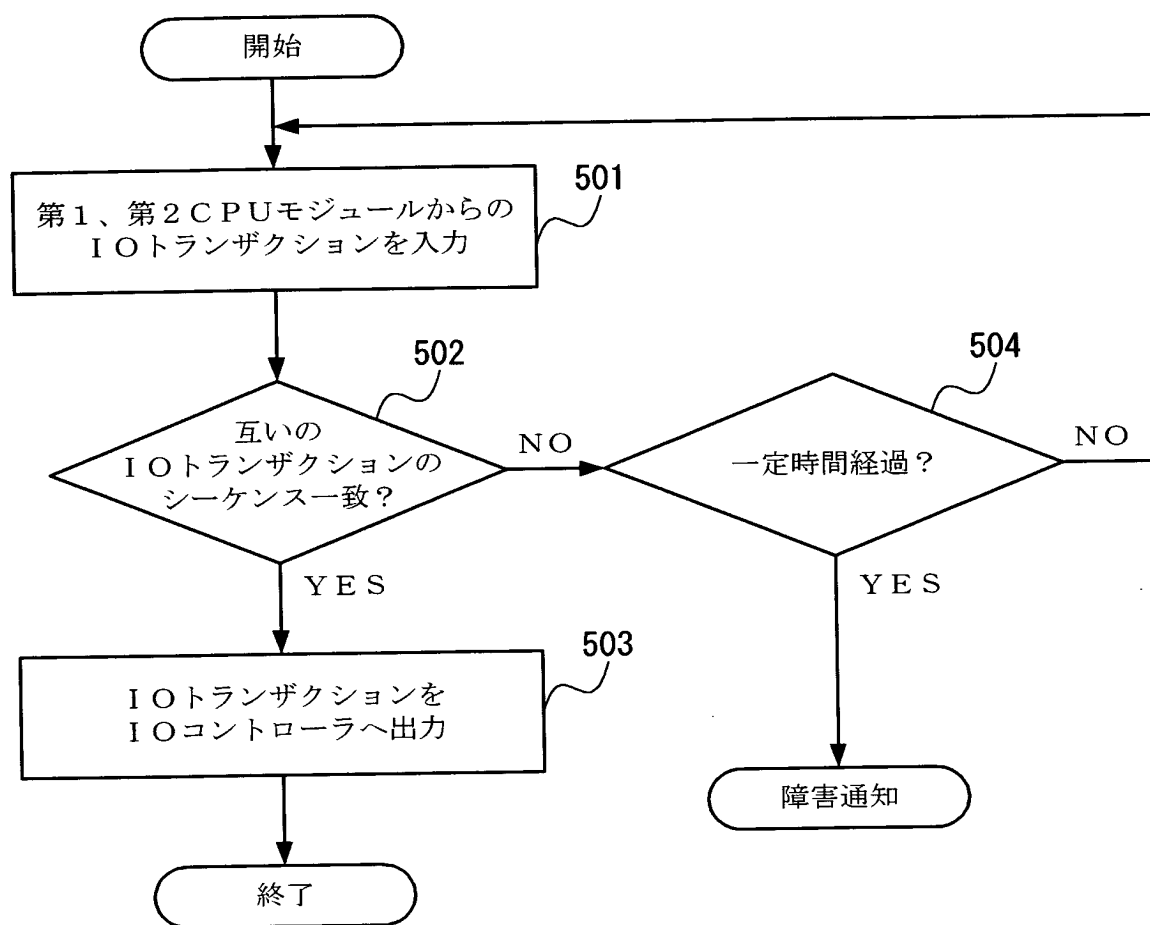
【図 3】



【図 4】



【図 5】



【書類名】 要約書

【要約】

【課題】 一定範囲のプロセッサの同期外れを許容することができるフォールトトレラントコンピュータ及びトランザクション同期制御プログラムを提案する。

【解決手段】 複数のCPUモジュール20、30で、同じ命令列をクロック同期させて処理するロックステップ方式のフォールトトレラントコンピュータ10であって、複数のCPUモジュール20、30から発行されるIOトランザクションについて、IOモジュール50、60のデバイスコントローラ毎に、一定の時間内において互いの一致をチェックし、一致が得られたIOトランザクションを同期外れではないとみなす。

【選択図】 図1

認定・付加情報

特許出願の番号	特願 2 0 0 2 - 2 5 5 3 8 8
受付番号	5 0 2 0 1 3 0 1 5 4 4
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 4 年 9 月 2 日

< 認定情報・付加情報 >

【提出日】 平成14年 8月30日

次頁無

特願 2002-255388

出 願 人 履 歴 情 報

識別番号

[000168285]

1. 変更年月日

1990年 8月 9日

[変更理由]

新規登録

住 所

山梨県甲府市大津町1088-3

氏 名

甲府日本電気株式会社

2. 変更年月日

2002年 7月30日

[変更理由]

名称変更

住所変更

住 所

山梨県甲府市大津町1088-3

氏 名

エヌイーシーコンピュータテクノ株式会社